COVER SHEET

Maire, Frederic and Rakotonirainy, Andry (2006) Analysis of Driving Session Videos by Reverse Temporal Order Processing. In *Proceedings Third International Conference on Computer Graphics, Imaging and Visualization (CGIV 2006)*, pages pp. 255-260, 26-28 July 2006, Sydney, Australia.

Analysis of Driving Session Videos by Reverse Temporal Order Processing

Frederic Maire[1,2] and Andry Rakotonirainy[3]
NICTA[1], 300 Adelaide Street, Brisbane, QLD 4000, Australia
School of SEDC[2] – Faculty of IT, Queensland University of Technology, Brisbane 4000, Australia.
Centre of Accident Research and Road Safety Queensland[3]– Faculty of Health,
Queensland University of Technology, Carseldine 4034 Australia
Frederic.Maire@nicta.com.au, {r.andry, f.maire}@qut.edu.au

Abstract

Technology based driver training is still in its infancy. There is a strong need for improved and integrated computer-based screening tools in order to facilitate objective and reliable driver training assessments. This paper describes a system that analyses videos of driving sessions collected by on-board web-cameras. The system detects and tracks lane markings in order to estimate the relative position of the vehicle with respect to its lane. The system is computationally efficient as it exploits fully the off-line nature of the data. The analysis of the video recording is performed in reverse temporal order. The benefits of this approach compared to the forward analysis traditionally used are an improved robustness and a lower computational cost.

Keywords--- **Video processing, Driver training, RANSAC, Reverse temporal order.**

1. Introduction

The over-representation of young and inexperienced drivers in road crashes is widely acknowledged as one of the most persistent road safety problems. Death in motor vehicle crashes is the most common fatal injury in young people aged 15-24 years in Australia [7] In 2002, almost one-third of drivers killed in motor vehicle crashes on Australian roads were aged between 17 and 25 years despite this age group representing only 12% of the population [8] Furthermore, it is believed that 31.7% of the contributing factors to the crashes of young people can be attributed to inexperience, with inattention, illegal manoeuvres, alcohol and failing to give way or stop, making up the remainder of the top five contributing circumstances [9] The accident liability of new drivers drops sharply over the first 12 months or so after passing their test and continues to fall as more experience is gained. This suggests that driving experience rather than age at licensing may be the primary reason for the higher crash risk for novice drivers.

Almost 95% of road crashes are attributable to driving error [6] Thus, driver training remains an important road safety intervention to improve driving performance and abilities, particularly with young people. Most driver training programs rely on the subjective evaluation of the instructor to evaluate trainee's skill. Technology based driver training is still in its infancy. There is a need for improved and integrated computer-based screening tools in order to facilitate objective and reliable driver training assessments.

Lane keeping is a fundamental skill that inexperienced drivers need to acquire. Lane departure is often a benign behaviour which could be simply corrected with a slight steering movement. Failure to correct lane departure could lead to run-off-road or sideswipe crashes. About 20% of all crashes occur in these circumstances and road crashes cost Australia $17 billion each year.

The integration of lane departure system in a driving training program to assess driving behaviour has not been comprehensively used to date. The work presented in this paper is part of a larger research program which addresses the need of designing new technology based driver assessment tool with road safety benefits. The system that we are aiming to build will provide objective and detailed feedback on a driver's behaviour using recorded data concerning their control of the vehicle, physical motions, attentiveness, assertiveness, how they respond to hazards and perceptions of the surrounding traffic environment. The computer based tool will help driver instructors to objectively and accurately evaluate driver control of the vehicle. It will extend existing driver training programs to reduce post-licensing crash risk by responding to the call for driver training to focus on perceptual skill deficiencies that have been shown to be associated with high collision rates of novice drivers.

This work expands on the functionality already provided by a driver training tool named VigilVanguard

Proceedings of the International Conference on Computer Graphics, Imaging and Visualisation (CGIV'06)
0-7695-2606-3/06 $20.00 © 2006 IEEE

[4] VigilVanguard is a portable device which records and measures unobtrusively and continuously vehicle dynamics during a naturalistic driving session. VigilVanguard uses GPS, advanced sensors and video cameras to measure and record several aspects of the driving behaviour, including speed and passenger comfort (braking, cornering forces). The data logged include speed, acceleration, following distance, GPS, and four web-cams views. VigilVanguard monitors driving performance during training on actual driving routes, providing a powerful on-road training tool and an objective assessment of drivers' abilities and reactions, leading to safer driving.

This paper presents a detection and tracking system for longitudinal lane markings to estimate the vehicle's position relative its lane. As VigilVanguard does not automatically analyse the video files; we have developed a prototype system to automatically analyse the video recordings to estimate the vehicle's position relative to its lane with the view of improving the accuracy of the instructor assessment. Note that the image processing is not done in real-time, but after the driving sessions. We have exploited the fact that we can analyse the video in both directions temporally to produce a robust video processing system. Our approach is computationally efficient and use neither digital mapping nor GPS.

Recording and analysing the driving practices of trainees will help the driver instructor to *(i)* objectively monitor the progress of the trainee during the entire training program and *(ii)* accurately debrief the trainee about their driving behaviour by showing driving errors through software featuring video footages. Our system will provide functions to *(i)* classify automatically the safe and unsafe driver behaviour (*ii*) compare manoeuvres, and *(iii)* navigate through different sequences of footage. Benefits have been shown to be gained by using recorded videos of novice driver's vehicle control strategies as feedback to teach drivers methods to improve their vehicle control skills [11]

In this paper, we focus on the video processing of driving sessions. More specifically, on the detection and tracking of lane markings. In Section 2, we explain how lane markings are detected and tracked. In Section 3, we present experimental results.

2. System architecture

2.1 Lane marking detection

Most lane marking detectors are too sensitive to lighting conditions [5] To address this problem, we use a variation of the algorithm presented in [1] to detect lane-marking centres. We scan each row of the grey image looking for pairs of points (A,B) having respectively a positive and negative gradient such that both the positive and the negative gradients have a magnitude larger than a minimum threshold. Moreover we require that the pixels between the two points A and B have an intensity larger than the sum of the intensity at A plus one half of the intensity gradient at A. Another constraint on (A,B) is that the distance between these two points lie in a predefined interval (prior knowledge that the width of lane-markings will be between 2 and 20 pixels). As can be seen in Figures 4, 5 and 6, this lane centre detection system is robust to bad lighting conditions; the trees' shadows do not affect the performance of the system.

2.3 Road model

As in [2] , we assume that the road is planar and that the lane markings can be approximated by a polynomial of degree d; $y_R = \sum_{i=0}^{d} a_i \, x_R^i$, where x_R denotes the coordinate with respect to the longitudinal axis of the road, and y_R is the coordinate with respect to an axis perpendicular to the road. The transformation between the road planer (x_R, y_R) and the image plane (x_I, y_I) is $x_I = c_x \dfrac{1}{x_R}$ and $y_I = c_y \dfrac{y_R}{x_R}$, where c_x and c_y are camera calibration parameters. The origin of the image coordinate system is taken on the horizon line. The projection of the polynomial curve $y_R = \sum_{i=0}^{d} a_i \, x_R^i$ is a *hyperbolic polynomial* curve of the form $y_I = \sum_{i=0}^{d} b_i \, x_I^{1-i}$. The assumption of a 2D flat road neglects the vertical curvature of the road. The main advantage is to allow estimation of the road shape from only one camera [2] For most vehicle control applications, the shape of the road has to be estimated only to a distance of 10 to 40 metres. This is why the flat road model is an acceptable solution in practice.

2.4 Tracking of the lane-markings

We analyse the video frames in reverse temporal order (play the tape backward). This choice is motivated by the following observations;

- When the tape is played backward, the new road features appear large and close to the car, whereas when the tape is played forward the new features appear small and near the horizon line. In other words, the new features are larger and clearer when the tape is played backward.
- When a new lane marking appears, it can be perfectly approximated with a straight line segment (hyperbolic polynomial models of degree 1).

For robust fitting of models, we use the RANSAC algorithm that was introduced by Fischler and Bolles in 1981 [10] The structure of the RANSAC algorithm is simple but powerful. First, samples are drawn uniformly and at random from the input data set. Each point has the same probability of selection (uniform point sampling). For each sample, a model hypothesis is constructed by computing the model parameters using the sample data. In the next step, the quality of the hypothetical models is evaluated on the full data set. The quality of the model is assessed by counting the number of inliers (data points which agree with the model within an error tolerance). The hypothesis which gets the most support from the data set gives the best estimate. The input data may support several distinct models. In this case, the model parameters for the first model are estimated, the data points supporting the model are removed from the input data and the algorithm is simply repeated with the remainder of the data set to find the next best model.

Below is the high level pseudo-code of the main loop of our program to compute for each frame F_t the set L_t of hyperbolic polynomials of degree d that represent the lane markings at time t.

```
LT+1 = empty set
for t = T down to 1  // T frames processed
   Ct = detect lane marking centres for frame Ft
   Lt = track Lt+1
   St = detect starts of new lane markings
   // St set of hyperbolic polynomials of degree 1
   Lt = merge St and Lt
end
```

The detection of the lane marking centres is performed as described in Section 2.1. For tracking the models of L_{t+1} we consider each model h (hyperbolic polynomial) in turn. For each model we determine C_t^h the subset of C_t of points within a certain distance of h. In our prototype, this is a hard coded constant (larger than the inlier threshold), but ideally this distance threshold should be determined by a model of the car dynamics. The subset C_t^h should be larger when the car is turning or travelling fast. RANSAC is run on C_t^h to track h. If the number of inliers for the revised model is too low, we consider that the lane marking has disappeared. If the model is maintained the points of C_t^h are removed from C_t before the computation of S_t. The starts of new lane markings are computed by examining only the bottom third of the image.

The detection of beginning of new lane-markings is achieved by running RANSAC on the bottom third of the image. This save computation time. The models sought for S_t are linear models. The models in S_t are then merged into L_t.

Notice that the models of S_t are hyperbolic polynomials of degree 1 which become hyperbolic polynomials of degree d when they join L_t (the coefficients $b_2,\ldots,b_d$ are simply set to zero.

The reverse temporal order processing saves some computation because we have only to analyse the bottom of the image to detect new lane markings, whereas a forward temporal order processing would have to analyse the whole part of the image below the horizon. An initial version of the system based on forward temporal order processing could only detect the new lane markings when it was run on the whole part of the image below the horizon. The other problem was that non-linear hyperbolic polynomials had to be used. To sum up, the computational gains of the reverse temporal order processing are due to the smaller region examined for new features and the fewer parameters needed for the models of the new lines. The reverse temporal order approach is more robust because it focuses on larger new features.

Once the lane markings are determined, the relative position of the car with respect to its lane can be estimated by computing the intersection of the hyperbolic polynomial curve with a fixed horizontal line in the image. In practice, this amounts to plugging the some value for x_I in $y_I = \sum_{i=0}^{d} b_i\, x_I^{1-i}$.

3 Experimental Results

The current prototype is implemented in Matlab. Figures 1, 4 and 7 are typical images obtained after converting the RGB video captured frames to grey images. The sensor device that can be seen on the hood of the car is a *following distance* sensor. In these experiments, we have set $d = 2$. That is, the lane markings image coordinates are assumed to be of the form $y_I = b_0\, x_I^1 + b_1 + b_2\, x_I^{-1}$.

Figure 1 Frame 07627

Figures 2, 5 and 8 show the lane marking centres detected. The white pixels indicate the centres of the lane markings. The surrounding grey intervals correspond to detected lane markings. Although many false positives appear in the images, the RANSAC algorithm consistently detects the lane markings.

Figure 2 Marking Centres of Frame 07627

The bottom left region of Figure 2 contains a significant amount of glare, generating many false positive lame marking centres. However, because we keep only hypothesised curves that have a large number of inliers, the system can cope with this noise as demonstrated in Figure 3.

Figure 3 Detected Markings for Frame 07627

The line labelled in red with "1" (the one on the right) has the higher confidence. Notice that the line labelled in red with "2" (the one on the left) does not fit perfectly the road near the horizon. This problem could probably be addressed by searching for couples lines with RANSAC. Indeed, it can be shown that the coefficients of the two lines should be equal except for the coefficient b_1. We have not yet tried this idea.

Figure 4 Frame 07815

The left-hand side of the road in Figure 4 is in the shade whereas the right-hand side is in the sun. Despite these poor lightings conditions, enough lane marking centres are detected on the left (see Figure 5). The road boundaries are properly detected as can be seen in Figure 6.

Figure 5 Marking Centres of Frame 07815

Figure 6 Detected Markings for Frame 07815

Figure 7 shows a complex situation with arrow markings on the ground. Many false lane marking centres are detected in this image. In particular the edge features of the cars are easily confused with lane marking centres.

Figure 7 Frame 13932

Although the system does not break down for this particular image, on other images the system will fail to detect the lane markings (confidence too low). An analysis of the ground colour and texture could eliminate some of the false positives and make the system even more robust.

Figure 8 Marking Centres of Frame 13932

In Figure 9, the system has detected three lane marking lines. The arrow was classified as a start of a new lane marking. More post-processing is needed to resolve this type of situations. In the current prototype, an ad hoc solution relying on the average observed lane width is used.

Figure 9 Detected Markings for Frame 13932

4 Conclusions

We have introduced a new method for lane marking detection and tracking based on the reverse temporal order analysis of video recordings. This technique is appropriate because the analysis of the video recordings does not have to be on-line or in real-time. This original approach simplifies the tracking as the new features on the road appear close to the vehicle. Moreover as the new features are straight line segments, a linear model can be used for their detection. The tracking is performed with more sophisticated curve models, namely hyperbolic polynomial curves. The system has its limitations. It loses track of the lanes at some road intersections. However, the information extracted from the video is sufficient to allow the automatic analysis of driver performance. Lane keeping behaviour can be assessed directly with the current system. Moreover, with the combinations of data from other sensors, it becomes possible to analyse automatically more complex manoeuvres like overtaking.

This paper presents the preliminary results of a larger project on building a computer based driver training tool. The future system will integrate knowledge about driver attitudes, self-reported behaviour, and actual driving performance on the road as measured with a set of Intelligent Transport Systems (ITS).

Acknowledgements

This project is partially supported by a FAST (French Australian Science and Technology) grant and the Motor Accident Insurance Commission (MAIC)-Australia.

References

[1] Ieng S-S and Tarel J-P. and Labayrade R. "On the Design of a Single Lane-Markings Detector Regardless the On-board Camera's Position", in Proceedings of IEEE Intelligent Vehicle Symposium (IV'2003),June 9-11 2003, Colombus, OH, USA.

[2] Ieng S-S "Methodes robustes pour la detection et le suivi des marquages”, These de doctorat de l’universite de Paris 6, Novembre 2004.

[3] Nuria Oliver, and Alex Pentland, "Graphical Models for Driver Behavior Recognition in a SmartCar", Proceedings of IEEE Intl. Conference on Intelligent Vehicles 2000 Detroit. Michigan. October 00.

[4] Vigil Systems, url: http://www.vigil-systems.com/

[5] Apostoloff, N. and Zelinsky A., Vision In and Out of Vehicles: Integrated Driver and Road Scene Monitoring, The International Journal of Robotic Research, Vol 23, No 4-5, April-May 2004, pp 513-538.

[6] Shinar, D. (1978). *Psychology on the Road: the Human Factor in Traffic Safety.* John Wiley & Sons, New York.

[7] Australian Bureau of Statistics. (2003). *Causes of Death, Australia 2002.* ABS Catalogue No 3303.0. Canberra: Author.

[8] Australian Transport Safety Bureau. (2003). *Road crash data and rates. Australian States and Territories 1925 to 2002.* Canberra: Department of Transport and Regional Services

[9] Queensland Transport (2002) Road Traffic Crashes in Queensland 2002. Qld Government - Qld Transport

[10] M. A. Fishler, R. C. Bolles. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. Comm. of the ACM, Vol 24, pp 381-395, 1981

[11] Chapman, P., Underwood, G., & Roberts, K. (2002). Visual search patterns in trained and untrained novice drivers. *Transportation Research Part F: Traffic Psychology & Behaviour, 5,* 157-167.